DOCUMENT RESUME

ED 236 203 TM 830 712

AUTHOR Fuchs, Douglas; Fuchs, Lynn S.
TITLE The Importance of Scorer Bias to Handicapped Preschoolers' Stronger Performance with Familiar Examiners.
INSTITUTION Minnesota Univ., Minneapolis. Inst. for Research on Learning Disabilities.
SPONS AGENCY Office of Special Education and Rehabilitative Services (ED), Washington, DC.
REPORT NO IRLD-RR-135
PUB DATE Jul 83
CONTRACT 300-80-0622
NOTE 22p.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Examiners; Experimenter Characteristics; *Language Handicaps; *Performance Factors; Preschool Education; Scoring; *Speech Evaluation; Test Anxiety; *Test Bias; Testing; Testing Problems; Videotape Recordings
IDENTIFIERS *Familiarity; Testing Conditions

ABSTRACT

Prior research indicates that language-handicapped children obtain higher test scores when tested by personally familiar examiners than when tested by personally unfamiliar examiners. The present investigation inquired whether this finding is due to examinees' actual differential performance across the two examiner conditions, or whether it is the result of testers' biased scoring of similar examinee performances. To make this determination, videotaped testing sessions, in which language-handicapped preschoolers were awarded higher scores by familiar examiners than by unfamiliar examiners, were shown to two certified speech clinicians who were blind to all purposes of the study. These individuals rated each examinee's performance in the familiar and unfamiliar examiner condition. Results indicated that the videotape raters, as the examiners, gave higher scores to examinees' performance in the familiar condition, corroborating the notion that language-impaired children actually perform more strongly with a familiar examiner. (Author)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED236203

University of Minnesota

Research Report No. 135

THE IMPORTANCE OF SCORER BIAS TO HANDICAPPED PRESCHOOLERS' STRONGER PERFORMANCE WITH FAMILIAR EXAMINERS

Douglas Fuchs and Lynn S. Fuchs

SCOPE OF INTEREST NOTICE
The ERIC Facility has assigned this document for processing to: TM EC
In our judgement, this document is also of interest to the clearinghouses noted to the right. Indexing should reflect their special points of view. PS

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
J. Ysseldyke
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

iRLD

Institute for Research on Learning Disabilities

TM 830 712

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
X This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

IRLD

Director: James E. Ysseldyke

The Institute for Research on Learning Disabilities is supported by a contract (300-80-0622) with Special Education Programs, Department of Education. Institute investigators are conducting research on the assessment/decision-making/intervention process as it relates to learning disabled students.

During 1980-1983, Institute research focuses on four major areas:

- Referral
- Identification/Classification
- Intervention Planning and Progress Evaluation
- Outcome Evaluation

Additional information on the Institute's research objectives and activities may be obtained by writing to the Editor at the Institute (see Publications list for address).

The research reported herein was conducted under government sponsorship. Contractors are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent the official position of Special Education Programs.

Research Report 135

THE IMPORTANCE OF SCORER BIAS TO HANDICAPPED PRESCHOOLERS' STRONGER PERFORMANCE WITH FAMILIAR EXAMINERS

Douglas Fuchs

Department of Education

Clark University

Lynn S. Fuchs

Wheelock College

July, 1983

Abstract

Prior research indicates that language-handicapped children obtain higher test scores when tested by personally familiar examiners than when tested by personally unfamiliar examiners. The present investigation inquired whether this finding is due to examinees' actual differential performance across the two examiner conditions, or whether it is the result of testers' biased scoring of similar examinee performances. To make this determination, videotaped testing sessions, in which language-handicapped preschoolers were awarded higher scores by familiar examiners than by unfamiliar examiners, were shown to two certified speech clinicians who were blind to all purposes of the study. These individuals rated each examinee's performance in the familiar and unfamiliar examiner condition. Results indicated that the videotape raters, as the examiners, gave higher scores to examinees' performance in the familiar condition, corroborating the notion that language-impaired children actually perform more strongly with a familiar examiner.

The Importance of Scorer Bias to Handicapped Preschoolers' Stronger Performance with Familiar Examiners

During the past four years, Fuchs and associates conducted a program of research into the effects of examiner familiarity on the performance of language-impaired children. Findings indicated that these children performed more strongly when they were assessed by familiar testers than when they were tested by strange examiners. More specifically, this differential performance was obtained (a) when testers were inexperienced and also when they were professional speech clinicians (Fuchs, Fuchs, Dailey, & Power, 1983), (b) across studies employing experimentally-induced (e.g., Fuchs, Fuchs, Power, & Dailey, in press) and long-term acquaintanceship (e.g., Fuchs, Fuchs, Garwick, & Featherstone, 1983) definitions of examiner familiarity, (c) over various levels of item difficulty and response modes (Fuchs, Featherstone, Garwick, & Fuchs, in press), and (d) across preschool and school-age language-impaired children (Fuchs, Fuchs, Power, & Dailey, 1983). Finally, this program of research demonstrated that the personal unfamiliarity of a tester not only discourages language-impaired children's optimal, absolute performance but also selectively depresses their performance relative to nonhandicapped children (Fuchs, Fuchs, Power, & Dailey, 1983).

Therefore, it appears that the effect of a tester's professional unfamiliarity prevails across a range of situations and that the use of unfamiliar examiners represents systematic bias against and threatens the validity of the test performance of certain handicapped children. The salience of these findings is underscored by the facts

that children typically are assessed by strange testers, most test manuals do not prescribe pretest contact between examiners and examinees (cf. Fuchs, Fuchs, Dailey, & Power, 1983), and test results are used pervasively for making decisions about educational programs and student classification.

Given the potentially negative, far-reaching implications that examiner unfamiliarity has for educational practice, it seems important to explore how and why a tester's strangeness affects certain examinees' performance. As a beginning, it may be useful to recognize an assumption that has been made explicitly and repeatedly in the paper thus far (as well as in all of the pertinent research to date): namely, that language-impaired children perform more strongly in the familiar condition. It is possible, of course, that examiner familiarity does not affect the level of examinees' responding but rather influences the accuracy of testers' judgment and scoring. A large and enduring literature on rater bias supports this latter possibility (e.g., Guilford, 1936; Rosenthal, 1980). In an effort to become clearer about the nature of examiner familiarity effects, the present study explored the impact of personal familiarity on testers' accuracy of scoring.

Method

Subjects

Subjects were 22 (17 M, 5 F) Caucasian preschool children. The $\bar{X}$ and SD for their CA were 58.32 and 8.70 months, respectively. They came from predominantly middle class homes in Central Massachusetts, were moderately to profoundly language-impaired, and attended a public

special education preschool program. All subjects performed within the normal range on individually administered intelligence tests.

Examiners

Examiners were 22 Caucasian female graduate students at a state college, and employees of public and private schools in Central Massachusetts. (See Fuchs, Fuchs, Dailey, & Power, 1983, for examiner selection procedure.) Eleven examiners were early childhood educators (ECEs), who had an average 96.00 months (SD = 59.28) teaching experience. None had formal training or professional experience with either assessment or handicapped children. For this reason, they were conceptualized as the "inexperienced" examiner group.

The other 11 examiners were speech clinicians (SCs) who had been practicing professionals for an average 85.09 months (SD = 74.06). By virtue of their professional experience and formal training addressing both assessment and language-handicapped youngsters, the SCs were assigned "experienced" examiner status. The two examiner groups were similar with respect to the amount of their respective work experiences, $t(20) = .38$, ns.

Measure

The Preschool Language Scale, verbal expression scale (VE; Zimmerman, Steiner, & Pond, 1979) was employed. Zimmerman et al. reported split-half reliability coefficients ranging from .75 to .95, with a median of .88 on the total test. Using the Spearman-Brown Prophecy formula, reliability for the VE was estimated at .79.

Design

Children were assigned randomly to SC and ECE groups. There was no difference between the two groups with respect to the children's

CA, $\underline{t}(20) = .75$, ns, or sex, $\chi^2(1) = 0.00$, ns. Within examiner groups, each child was assigned randomly to two examiners, one with whom he or she became personally familiar and one to whom he or she remained a stranger. The study required each examiner to serve in both familiar and unfamiliar roles, thereby controlling for potentially confounding effects of tester personality. Each child was assessed twice during a period of three weeks, once by the familiar and once by the unfamiliar tester, within a crossover design: One-half of both ECE and SC examiners first tested familiar children, then unfamiliar children; the remaining examiners tested their examinees in reverse order. All testing occurred in the preschool's speech therapy room, a setting with which all children were familiar.

Procedure

Personal familiarity. Examiners' personal familiarity was induced experimentally by two procedures. Every tester was required to make a one-hour home visit. Examiners were told that there were two purposes for this visit: first, "to get to know the child and to permit the child to get to know you"; second, "to obtain information about the child from the mother." Accordingly, each tester was instructed to take materials with which to play with her future examinee and to administer to the child's mother a structured interview that briefly explored the child's general functioning and likes and dislikes. (Although scored and returned to the investigators, the interview data were not subjected to analyses; the only purpose for the interview was to acquaint the examiner with the child.)

The second strategy to induce personal familiarity was to require each tester to play with her "familiar" child for one hour immediately preceding the tester session. The play occurred both in and outside of the child's classroom. For this encounter, the tester provided the same materials with which she and the child had played during the home visit. The play outside of the child's room always followed the classroom interaction; the preschool encounters always followed the examiners' home visits. The lapse in time separating the home visit and testing ranged between two and eight days.

Training. ECE and SC examiners were trained separately to administer the VE scale. The ECEs received a total of five hours of instruction in two sessions. The SCs met for one session that lasted 2½ hours. A certified speech clinician conducted all training.

Videotaping. The students' test performance with familiar and unfamiliar examiners was videotaped with two AVC 3200 Sony video cameras on one-half inch videotape. The cameras, connected to a Sony 3600 recorder, were placed behind the examiner and examinee. With the aid of a special effects generator (SEG-1), a split screen was created displaying a frontal view of the upper torsos and heads of both participants. Examiners were informed of the recording; examinees were not.

Scoring. Examinees' performances on the VE, in both familiar and unfamiliar testing conditions, were scored using two procedures. First, investigators summed examiners' protocols that had been completed during testing, using a blind procedure so that investigators were unaware of examinees' names or testing conditions.

For the second scoring, two female certified speech clinicians, who did not know (a) any of the examiners or examinees, (b) the purpose of the study, or (c) the testing conditions they viewed, completed new protocols as they watched the videotaped testing sessions. These raters scored equal numbers of SC and ECE examiner testing sessions. One rater observed 45% and 55% of familiar and unfamiliar testing sessions, respectively, with the remaining sessions scored by the other rater. Interscorer agreement,[1] calcalated on 18% of the testing, ranged between .91 and .96. Later, the second set of protocols was summed by investigators using a blind procedure.

Results

A preliminary one between (SC vs. ECE), one within (personally familiar vs. unfamiliar) analysis of variance (ANOVA) was run on the VE scores (Winer, 1971). This ANOVA yielded one significant effect for personal familiarity vs. unfamiliarity, $F(1,20) = 3.56$, $p < .05$. Across experienced and inexperienced examiner conditions, subjects performed an average 4.11 points higher when tested by personally familiar examiners.

Next, for each child's performance, both in familiar and unfamiliar testing conditions, a difference score was calculated between the examiner's VE score and the videotape rater's VE score. These difference scores, indicative of the examiners' scoring accuracy, were entered into a one within (personally familiar vs. unfamiliar) ANOVA, which revealed no statistically significant difference in the scoring accuracy of examiners between familiar and unfamiliar testing conditions, $F(1,21) = .46$, ns. The average

differences between examiners' and videotaped raters' scores were 4.77 (SD = 5.83) and 3.61 (SD = 5.79), in the familiar and unfamiliar conditions, respectively.

Discussion

Results indicate that a nonsignificant disparity was generated by contrasting examiners' vs. videotape raters' scores in the familiar testing condition with those in the unfamiliar test setting. Thus, the videotape raters also obtained higher scores for the examinees' performance in the familiar condition. Because these raters knew nothing about the study's objectives or participants, the findings seem to support the notion that examinees actually performed differently across the two experimental conditions, rather than performing similarly and receiving differential scores by biased testers.

Although we found no evidence indicating that examiners' biased scoring was responsible for examinees' differential performance, we do not wish to imply that an examinee's performance is independent of tester behavior. A previous study (Fuchs, Zern, & Fuchs, in press-a; Fuchs, Zern, & Fuchs, in press-b) demonstrated an association between children's differential verbal production in familiar vs. unfamiliar examiner conditions and examiner behavior. In the familiar condition, examinees spoke longer, more often, and with greater syntactic and semantic complexity; familiar examiners (a) exercised more frequent and longer intervals of silence than unfamiliar examiners, (b) often used eye contact with examinees as a cue in deciding when to speak, whereas unfamiliar examiners rarely utilized this cue, (c) employed

largely directive language in contrast to unfamiliar examiners' speech that more frequently was participatory in nature, and (d) spoke for shorter durations than unfamiliar examiners.

Thus, this study and previous related investigations suggest that the situational factor, examiner familiarity, affects both examiner and examinee behavior in dramatic and educationally significant ways. Examiner trainers, test developers and publishers, professional groups that are responsible for establishing and monitoring testing standards, researchers, and users of test findings should consider more seriously the role of tester familiarity and, simultaneously, begin to question the possible importance of additional, unexplored situational factors in the test situation to children's performance.

References

Fuchs, D., Featherstone, N., Garwick, D. R., & Fuchs, L. S. Effects of examiner familiarity and task characteristics on speech and language-impaired children's test performance. Measurement and Evaluation in Guidance, in press.

Fuchs, D., Fuchs, L. S., Dailey, A. M., & Power, M. H. Effects of pretest contact with experienced and inexperienced examiners on handicapped children's performance (Research Report No. 110). Minneapolis: University of Minnesota, Institute for Research on Learning Disabilities, 1983.

Fuchs, D., Fuchs, L. S., Garwick, D. R., & Featherstone, N. Test performance of language-handicapped children with familiar and unfamiliar examiners. Journal of Psychology, 1983, 114, 37-46.

Fuchs, D., Fuchs, L. S., Power, M. H., & Dailey, A. M. Systematic bias in the assessment of handicapped children (Research Report No. 134). Minneapolis: University of Minnesota, Institute for Research on Learning Disabilities, 1983.

Fuchs, D., Zern, D. S., & Fuchs, L. S. A microanalysis of participant behavior in familiar and unfamiliar test conditions. Exceptional Children, in press. (a)

Fuchs, D., Zern, D. S., & Fuchs, L. S. Participants' verbal and nonverbal behavior in familiar and unfamiliar test conditions: An exploratory analysis. Diagnostique, in press. (b)

Guilford, J. P. Psychometric methods. New York: McGraw Hill, 1936.

Rosenthal, R. Experimenter effects in behavioral research (Enlarged paperback edition). New York: Irvington, 1980.

Thompson, R. H., White, K. R., & Morgan, D. P. Teacher-student interaction patterns in classrooms with mainstreamed mildly handicapped students. American Educational Research Journal, 1982, 19, 220-236.

Winer, B. J. Statistical principles in experimental design (2nd ed.). New York: McGraw-Hill, 1971.

Zimmerman, I. L., Steiner, V. G., & Pond, R. E. Preschool language scale. Columbus, OH: Charles E. Merrill, 1979.

Footnote

Douglas Fuchs and Lynn S. Fuchs also are affiliated with the Institute for Research on Learning Disabilities.

[1]Interobserver agreement was calculated using the following formula (Coulter cited in Thompson, White, & Morgan, 1982):

$$\text{Percentage Agreement} = \frac{\text{Agreements between Observer A and Observer B}}{\text{Agreements between A \& B + Disagreements between A \& B + Omissions by A + Omissions by B}}$$

PUBLICATIONS

Institute for Research on Learning Disabilities
University of Minnesota

The Institute is not funded for the distribution of its publications. Publications may be obtained for $4.00 each, a fee designed to cover printing and postage costs. Only checks and money orders payable to the University of Minnesota can be accepted. All orders must be prepaid. Requests should be directed to: Editor, IRLD, 350 Elliott Hall; 75 East River Road, University of Minnesota, Minneapolis, MN 55455.

The publications listed here are only those that have been prepared since 1982. For a complete, annotated list of all IRLD publications, write to the Editor.

Wesson, C., Mirkin, P., & Deno, S. Teachers' use of self instructional materials for learning procedures for developing and monitoring progress on IEP goals (Research Report No. 63). January, 1982.

Fuchs, L., Wesson, C., Tindal, G., Mirkin, P., & Deno, S. Instructional changes, student performance, and teacher preferences: The effects of specific measurement and evaluation procedures (Research Report No. 64). January, 1982.

Potter, M., & Mirkin, P. Instructional planning and implementation practices of elementary and secondary resource room teachers: Is there a difference? (Research Report No. 65). January, 1982.

Thurlow, M. L., & Ysseldyke, J. E. Teachers' beliefs about LD students (Research Report No. 66). January, 1982.

Graden, J., Thurlow, M. L., & Ysseldyke, J. E. Academic engaged time and its relationship to learning: A review of the literature (Monograph No. 17). January, 1982.

King, R., Wesson, C., & Deno, S. Direct and frequent measurement of student performance: Does it take too much time? (Research Report No. 67). February, 1982.

Greener, J. W., & Thurlow, M. L. Teacher opinions about professional education training programs (Research Report No. 68). March, 1982.

Algozzine, B., & Ysseldyke, J. Learning disabilities as a subset of school failure: The oversophistication of a concept (Research Report No. 69). March, 1982.

Fuchs, D., Zern, D. S., & Fuchs, L. S. A microanalysis of participant behavior in familiar and unfamiliar test conditions (Research Report No. 70). March, 1982.

Shinn, M. R., Ysseldyke, J., Deno, S., & Tindal, G. A comparison of psychometric and functional differences between students labeled learning disabled and low achieving (Research Report No. 71). March, 1982.

Thurlow, M. L. Graden, J., Greener, J. W., & Ysseldyke, J. E. Academic responding time for LD and non-LD students (Research Report No. 72). April, 1982.

Graden, J., Thurlow, M., & Ysseldyke, J. Instructional ecology and academic responding time for students at three levels of teacher-perceived behavioral competence (Research Report No. 73). April, 1982.

Algozzine, B., Ysseldyke, J., & Christenson, S. The influence of teachers' tolerances for specific kinds of behaviors on their ratings of a third grade student (Research Report No. 74). April, 1982.

Wesson, C., Deno, S., & Mirkin, P. Research on developing and monitoring progress on IEP goals: Current findings and implications for practice (Monograph No. 18). April, 1982.

Mirkin, P., Marston, D., & Deno, S. L. Direct and repeated measurement of academic skills: An alternative to traditional screening, referral, and identification of learning disabled students (Research Report No. 75). May, 1982.

Algozzine, B., Ysseldyke, J., Christenson, S., & Thurlow, M. Teachers' intervention choices for children exhibiting different behaviors in school (Research Report No. 76). June, 1982.

Tucker, J., Stevens, L. J., & Ysseldyke, J. E. Learning disabilities: The experts speak out (Research Report No. 77). June, 1982.

Thurlow, M. L., Ysseldyke, J. E., Graden, J., Greener, J. W., & Mecklenberg, C. Academic responding time for LD students receiving different levels of special education services (Research Report No. 78). June, 1982.

Graden, J. L., Thurlow, M. L., Ysseldyke, J. E., & Algozzine, B. Instructional ecology and academic responding time for students in different reading groups (Research Report No. 79). July, 1982.

Mirkin, P. K., & Potter, M. L. A survey of program planning and implementation practices of LD teachers (Research Report No. 80). July, 1982.

Fuchs, L. S., Fuchs, D., & Warren, L. M. Special education practice in evaluating student progress toward goals (Research Report No. 81). July, 1982.

Kuehnle, K., Deno, S. L., & Mirkin, P. K. Behavioral measurement of social adjustment: What behaviors? What setting? (Research Report No. 82). July, 1982.

Fuchs, D., Dailey, Ann Madsen, & Fuchs, L. S. Examiner familiarity and the relation between qualitative and quantitative indices of expressive language (Research Report No. 83). July, 1982.

Videen, J., Deno, S., & Marston, D. Correct word sequences: A valid indicator of proficiency in written expression (Research Report No. 84). July, 1982.

Potter, M. L. Application of a decision theory model to eligibility and classification decisions in special education (Research Report No. 85). July, 1982.

Greener, J. E., Thurlow, M. L., Graden, J. L., & Ysseldyke, J. E. The educational environment and students' responding times as a function of students' teacher-perceived academic competence (Research Report No. 86). August, 1982.

Deno, S., Marston, D., Mirkin, P., Lowry, L., Sindelar, P., & Jenkins, J. The use of standard tasks to measure achievement in reading, spelling, and written expression: A normative and developmental study (Research Report No. 87). August, 1982.

Skiba, R., Wesson, C., & Deno, S. L. The effects of training teachers in the use of formative evaluation in reading: An experimental-control comparison (Research Report No. 88). September, 1982.

Marston, D., Tindal, G., & Deno, S. L. Eligibility for learning disability services: A direct and repeated measurement approach (Research Report No. 89). September, 1982.

Thurlow, M. L., Ysseldyke, J. E., & Graden, J. L. LD students' active academic responding in regular and resource classrooms (Research Report No. 90). September, 1982.

Ysseldyke, J. E., Christenson, S., Pianta, R., Thurlow, M. L., & Algozzine, B. An analysis of current practice in referring students for psychoeducational evaluation: Implications for change (Research Report No. 91). October, 1982.

Ysseldyke, J. E., Algozzine, B., & Epps, S. A logical and empirical analysis of current practices in classifying students as handicapped (Research Report No. 92). October, 1982.

Tindal, G., Marston, D., Deno, S. L., & Germann, G. Curriculum differences in direct repeated measures of reading (Research Report No. 93). October, 1982.

Fuchs, L.S., Deno, S. L., & Marston, D. Use of aggregation to improve the reliability of simple direct measures of academic performance (Research Report No. 94). October, 1982.

Ysseldyke, J. E., Thurlow, M. L., Mecklenburg, C., & Graden, J. Observed changes in instruction and student responding as a function of referral and special education placement (Research Report No. 95). October, 1982.

Fuchs, L. S., Deno, S. L., & Mirkin, P. K. Effects of frequent curriculum-based measurement and evaluation on student achievement and knowledge of performance: An experimental study (Research Report No. 96). November, 1982.

Fuchs, L. S., Deno, S. L., & Mirkin, P. K. Direct and frequent measurement and evaluation: Effects on instruction and estimates of student progress (Research Report No. 97). November, 1982.

Tindal, G., Wesson, C., Germann, G., Deno, S. L., & Mirkin, P. K. The Pine County model for special education delivery: A data-based system (Monograph No. 19). November, 1982.

Epps, S., Ysseldyke, J. E., & Algozzine, B. An analysis of the conceptual framework underlying definitions of learning disabilities (Research Report No. 98). November, 1982.

Epps, S., Ysseldyke, J. E., & Algozzine, B. Public-policy implications of different definitions of learning disabilities (Research Report No. 99). November, 1982.

Ysseldyke, J. E., Thurlow, M. L., Graden, J. L., Wesson, C., Deno, S. L., & Algozzine, B. Generalizations from five years of research on assessment and decision making (Research Report No. 100). November, 1982.

Marston, D., & Deno, S. L. Measuring academic progress of students with learning difficulties: A comparison of the semi-logarithmic chart and equal interval graph paper (Research Report No. 101). November, 1982.

Beattie, S., Grise, P., & Algozzine, B. Effects of test modifications on minimum competency test performance of third grade learning disabled students (Research Report No. 102). December, 1982

Algozzine, B., Ysseldyke, J. E., & Christenson, S. An analysis of the incidence of special class placement: The masses are burgeoning (Research Report No. 103). December, 1982.

Marston, D., Tindal, G., & Deno, S. L. Predictive efficiency of direct, repeated measurement: An analysis of cost and accuracy in classification (Research Report No. 104). December, 1982.

Wesson, C., Deno, S., Mirkin, P., Sevcik, B., Skiba, R., King, R., Tindal, G., & Maruyama, G. Teaching structure and student achievement effects of curriculum-based measurement: A causal (structural) analysis (Research Report No. 105). December, 1982.

Mirkin, P. K., Fuchs, L. S., & Deno, S. L. (Eds.). Considerations for designing a continuous evaluation system: An integrative review (Monograph No. 20). December, 1982.

Marston, D., & Deno, S. L. Implementation of direct and repeated measurement in the school setting (Research Report No. 106). December, 1982.

Deno, S. L., King, R., Skiba, R., Sevcik, B., & Wesson, C. The structure of instruction rating scale (SIRS): Development and technical characteristics (Research Report No. 107). January, 1983.

Thurlow, M. L., Ysseldyke, J. E., & Casey, A. Criteria for identifying LD students: Definitional problems exemplified (Research Report No. 108). January, 1983.

Tindal, G., Marston, D., & Deno, S. L. The reliability of direct and repeated measurement (Research Report No. 109). February, 1983.

Fuchs, D., Fuchs, L. S., Dailey, A. M., & Power, M. H. Effects of pre-test contact with experienced and inexperienced examiners on handicapped children's performance (Research Report No. 110). February, 1983

King, R. P., Deno, S., Mirkin, P., & Wesson, C. The effects of training teachers in the use of formative evaluation in reading: An experimental-control comparison (Research Report No. 111). February, 1983.

Tindal, G., Deno, S. L., & Ysseldyke, J. E. Visual analysis of time series data: Factors of influence and level of reliability (Research Report No. 112). March, 1983.

Tindal, G, Shinn, M., Fuchs, L., Fuchs, D., Deno, S., & Germann, G. The technical adequacy of a basal reading series mastery test (Research Report No. 113). April, 1983.

Sevcik, B., Skiba, R., Tindal, G., King, R., Wesson, C., Mirkin, P., & Deno, S. Communication of IEP goals and student progress among parents, regular classroom teachers, and administrators using systematic formative evaluation (Research Report No. 114). April, 1983.

Wesson, C. Two student self-management techniques applied to data-based program modification (Research Report No. 115). April, 1983.

Wesson, C., Skiba, R., Sevcik, B., King, R., Tindal, G., Mirkin, P., & Deno, S. The impact of the structure of instruction and the use of technically adequate instructional data on reading improvement (Research Report No. 116). May, 1983.

Wesson, C. Teacher vs student selection of instructional activities (Research Report No. 117). May, 1983.

Tindal, G., & Deno, S. Factors influencing the agreement between visual and statistical analyses of time series data (Research Report No. 118). June, 1983.

Skiba, R. S. Classroom behavior management: A review of the literature (Monograph No. 21), June, 1983.

Graden, J. L., Thurlow, M. L., & Ysseldyke, J. E. When are students most academically engaged? Academic responding time in different instructional ecologies (Research Report No. 119). June, 1983.

Fuchs, L. S., Deno, S. L., & Roettger, A. The effect of alternative data-utilization rules on spelling achievement: An n of 1 study (Research Report No. 120). June, 1983.

Skiba, R., Sevcik, B., Wesson, C., King, R., & Deno, S. The non-effect of process-product variables in resource classrooms (Research Report No. 121). June, 1983.

Fuchs, L. Tindal, G., Shinn, M., Fuchs, D., Deno, S., & Germann, G. Technical adequacy of basal readers' mastery tests: The Ginn 720 series (Research Report No. 122). June, 1983.

Tindal, G., Germann, G., Marston, D., & Deno, S. The effectiveness of special education: A direct measurement approach (Research Report No. 123). June, 1983.

Sevcik, B., Skiba, R., Tindal, G., King, R., Wesson, C., Mirkin, P., & Deno, S. Curriculum-based measurement: Effects on instruction, teacher estimates of student progress, and student knowledge of performance (Research Report No. 124). July, 1983.

Skiba, R., Marston, D., Wesson, C., Sevcik, B., & Deno, S. L. Characteristics of the time-series data collected through curriculum-based reading measurement (Research Report No. 125). July, 1983.

Ysseldyke, J., Christenson, S., Graden, J., & Hill, D. Practical implications of research on referral and opportunity to learn (Monograph No. 22). July, 1983.

Marston, D., Deno, S., & Tindal, G. A comparison of standardized achievement tests and direct measurement techniques in measuring pupil progress (Research Report No. 126). July, 1983.

Fuchs, D., Fuchs, L. S., Tindal, G., & Deno, S. L. Variability of performance: A "signature" characteristic of learning disabled children? (Research Report No. 127). July, 1983.

Tindal, G., Fuchs, L., Fuchs, D., Shinn, M., Deno, S., & Germann, G. The technical adequacy of a basal series mastery test: The Scott-Foresman reading program (Research Report No. 128). July, 1983.

Fuchs, L. S., Fuchs, D., & Deno, S. L. The nature of inaccuracy among readability formulas (Research Report No. 129). July, 1983.

Fuchs, L., Tindal, G., Fuchs, D., Shinn, M., Deno, S., & Germann, G. The technical adequacy of a basal reading mastery test: The Holt basic reading series (Research Report No. 130). July, 1983.

Ysseldyke, J. E., Christenson, S., Algozzine, B., & Thurlow, M. L. Classroom teachers' attributions for students exhibiting different behaviors (Research Report No. 131). July, 1983.

Tindal, G., Germann, G., & Deno, S. Descriptive research on the Pine County norms: A compilation of findings (Research Report No. 132). July, 1983.

Skiba, R. J. The relationship between classroom management strategies and student misbehaviors (Research Report No. 133). July, 1983.

Fuchs, D., Fuchs, L. S., Power, M. H., & Dailey, A. M. Systematic bias in the assessment of handicapped children (Research Report No. 134). July, 1983.

Fuchs, D., & Fuchs, L. S. The importance of scorer bias to handicapped preschoolers' stronger performance with familiar examiners (Research Report No. 135). July, 1983.